Interval Fuzzy Bayesian Inference

Juan Miguel León-Rojas
University of Extremadura
Department of Mathematics
Escuela Politécnica
Avda. de la Universidad, s/n
10003, Cáceres, Extremadura, Spain
jmleon@unex.es

Montaña Morales Morgado
Regional Government of Extremadura
Department for Education and Employment
montmorgado@gmail.com

This is a preprint; it is published under Creative Commons Attribution ShareAlike (CC BY-SA) license. Update: Once peer-reviewed, it has been published as a chapter in: Soft Methodology and Random Information Systems. Volume 26 of the series Advances in Soft Computing pp 559-566. DOI: 10.1007/978-3-540-44465-7_69. Print ISBN: 978-3-540-22264-4. Online ISBN: 978-3-540-44465-7. Series ISSN: 1867-5662. Publisher: Springer Berlin Heidelberg. URL: `http://link.springer.com/chapter/10.1007/978-3-540-44465-7_69` (© 2004).

UNESCO nomenclature: 1209.04 Decision procedures and theory; 1209.13 Techniques of statistical inference; 2201.05 Noise.

The aim of this paper is to show a way of applying Bayesian inference on interval fuzzy data assuming that all the time we work with interval probabilities. The proposal is illustrated with an example in Health Care, specifically on inferring population annoyance level caused by noise exposure.

0.1 Introduction

In this paper we introduce a way of applying Bayesian inference in the presence of linguistic imprecision at the same time in evidence, priors and marginals. Linguistic imprecision is represented using interval fuzzy and interval probability assignments. Our ideas are illustrated with an example in Health Care, specifically on inferring population annoyance level caused by noise exposure. Expressions such as "In the light of these results we can say that for a sound level of approximately six or seven, then, with a probability, approximately high or perhaps very high, the best alternative is that people living at that place is moderate annoyed because of acoustic contamination" are allowed.

0.2 Imprecise probabilities model

In the presence of imprecision, instead of a single probability distribution, we can use a class $\mathcal{M}$ of priors, so that each prior in $\mathcal{M}$ is updated by Bayes' rule, producing a class of posteriors. In the theory of *imprecise probabilities* the lower and upper probabilities of an event or hypothesis H, denoted by $\underline{P}(H)$ and $\overline{P}(H)$, are defined by $\underline{P}(H) = \inf\{P(H) \mid P \in \mathcal{M}\}$ and $\overline{P}(H) = \sup\{P(H) \mid P \in \mathcal{M}\}$. This theory denies that any precise probability model is reasonable. Statistical conclusions will be expressed in terms of posterior lower and upper probabilities or expectations, and these imprecise probabilities should not be regarded as lower and upper bounds for some unknown precise probability [1].

Let Θ be the parameter space and $A \subseteq \Theta$. The lower and upper posteriors of A are obtained minimizing and maximizing, respectively, the posterior probability

$$\pi(A|\mathbf{x}) = \int_A \pi^*(\theta|\mathbf{x})d\theta \Big/ \int_\Theta \pi^*(\theta|\mathbf{x})d\theta \tag{1}$$

in the lower/upper class of unnormalized posteriors $\pi^*(\ |\mathbf{x})$, which we assume being represented as an interval of measures [2]. In order to minimize $\pi(A|\mathbf{x})$, we use (Cf. [1])

$$\pi^*(\theta|\mathbf{x}) = l^*(\theta|\mathbf{x})\chi_{\theta\in A} + u^*(\theta|\mathbf{x})\chi_{\theta\in\Theta\setminus A} \tag{2}$$

in (1), where χ_P denotes the characteristic function of predicate P, i.e. $\chi_P = 1$ if P is true, and $\chi_P = 0$ if P is false. Thus, we assign the minimum possible mass to A and the maximum possible mass to the complement of A in Θ. The lower posterior is

$$\underline{P}\,(A|\mathbf{x}) = \int_A l^*(\theta|\mathbf{x})d\theta \Big/ \left(\int_A l^*(\theta|\mathbf{x})d\theta + \int_{\Theta\setminus A} u^*(\theta|\mathbf{x})d\theta\right) \tag{3}$$

In order to maximize $\pi(A|\mathbf{x})$, we only have to swap l^* with u^* in (2).

0.3 First Ideas on Fuzzy Bayesian Inference

A linguistic variable whose values are words or sentences may be defined by the quadruplet [3] $\langle x_{\text{name}}, \mathcal{L}(\mathcal{E}), \mathcal{E}, M\rangle$, where x_{name} is the name of the linguistic variable; $\mathcal{L}(\mathcal{E})$ is the *term-set* or *reference-set* of x_{name}, i. e., the *finite* set of linguistic values that x_{name} can take, which elements we denote by e_L; $\mathcal{E}$ is the *universe of discourse* or physical domain associated with x_{name}; M is a semantic function that associates a fuzzy meaning to each linguistic value $e_L \in \mathcal{L}(\mathcal{E})$, i. e., it is an injective mapping from $\mathcal{L}(\mathcal{E})$ to $\mathfrak{F}(\mathcal{E})$, the set of fuzzy subsets of $\mathcal{E}$. As it is more and more usual, we denote $M(e_L)$ by e_L. Remember that the *support* and the *core* of $A \in \mathfrak{F}(\mathcal{E})$ are $\text{supp}(A) = \{x \in\mathcal{E}\colon A(x) > 0\}$ and $\text{core}(A) = \{x \in\mathcal{E}\colon A(x) = 1\}$.

First works on fuzzy Bayesian inference come from safety project studies in structural reliability researches [4]. Given a linguistic value e_L of the evidence, and a set of exhaustive and mutually exclusive hypotheses H_j $(j = 1, \ldots, m)$, we can compute the likelihood $p(e_L|H_j)$ by

$$p(e_L|H_j) = \int_{e\in e_L} \mu_{e_L}(e) f(e|H_j)de \tag{4}$$

where $f(e|H_j)$ is the likelihood density function evaluated at e given the hypothesis H_j. We can compute the posterior probability by

$$p(H_j|e_L) = p(H_j)p(e_L|H_j) \Big/ \textstyle\sum_{k=1}^m \left(p(H_k)p(e_L|H_k)\right) \tag{5}$$

One problem is the determination of $f(e|H_j)$ (usually approximated as Gaussian or Weibull). On the other hand, once $f(e|H_j)$ has been determined, another problem is the computation of the integral itself. In order to save us from these problems, Yang [5] proposes to deduce $f(e|H_j)$ from $p(e_L|H_j)$:

$$f(e|H_j) = c \sum_{e_L\in\mathcal{L}(E)} \frac{\mu_{e_L}(e)}{W(e_L)} p(e_L|H_j); \quad \text{with } c = W(e_L) \Big/ \int_{e\in\text{supp}e_L} \mu_{e_L}(e)de$$

Lastly, we can compute the posterior probability by the Bayes' rule,

$$p(H_j|e) = f(e|H_j)p(H_j) \Big/ \textstyle\sum_{j=1}^m f(e|H_j)p(H_j) \tag{6}$$

The use of c means that the size $W(e_L)$ of the range covered by any linguistic value e_L, divided by the area of its membership function μ_{e_L}, is the same for each linguistic value e_L, and is equal to c.

0.4 *Interval Fuzzy Bayesian Inference (IFBI)*

Our proposal does not depend on any constant and it works with interval fuzzy and interval probability assignments. Let e_L be a fuzzy value for the evidence. Let $\Theta = \{H_1, \ldots, H_m\}$ be a complete set of hypotheses, i.e. a finite set of exhaustive and mutually exclusive hypotheses. We define the likelihood function for any continuous value $e \in \mathcal{E}$ of the evidence (given an hypothesis H_j), from the likelihoods of the fuzzy values e_L of the evidence (given the same hypothesis H_j):

$$L(H_j|e) \propto \sum\nolimits_{e_L \in \mathcal{L}(\mathcal{E})} e_L(e)p(e_L|H_j) \tag{7}$$

We extend (7) to fuzzy values $E \in \mathfrak{F}(\mathcal{E})$ of the evidence by

$$\widetilde{L}(H_j|E) \propto \sum\nolimits_{e_L \in \mathcal{L}(\mathcal{E})} \widetilde{e_L}(E)p(e_L|H_j) \tag{8}$$

where $\widetilde{e_L} : \mathfrak{F}(\mathcal{E}) \to \mathfrak{F}([0,1])$ extends $e_L : \mathcal{E} \to [0,1]$ according to Zadeh's extension principle (Cf. Footnote 1).

Let n, m be positive integer numbers. Let $\mathbb{I}^n\mathbb{R}$ be the set of all intervals of dimension n of real numbers. Given $D \subseteq \mathbb{R}^m$, we denote by $\mathbb{I}^m D$ the set of all intervals of dimension m included in D. Given $f : D \to \mathbb{R}^n$, an *inclusion function* for f, is any function $\Box f : \mathbb{I}^m D \to \mathbb{I}^n\mathbb{R}$, such that, $\forall J \in \mathbb{I}^m D$, $f(J) \subseteq \Box f(J)$. If f is monotonic increasing, such as e^x, $\ln x$, or $\sqrt{x}$, then an inclusion function is $\Box f([a_0, a_1]) = [f(a_0), f(a_1)]$. If f is monotonic decreasing, then an inclusion function is $\Box f([a_0, a_1]) = [f(a_1), f(a_0)]$. The case of non monotonic functions is not difficult if we know the intervals where the function is monotonic. For example, in the illustrative example below, we use *triangular fuzzy numbers* (tfn), because of their ease of computation. The membership function of a tfn is characterised by the lower, modal and upper values, i.e. the vertices of the triangle:

$$\mu(x; a, b, c) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{c-x}{c-b}\right), 0\right) \tag{9}$$

The natural extension of a triangular fuzzy number $T(x; a, b, c)$ is

$$\begin{aligned} \Box T([x]; a, b, c) = {} & [T(x_0; a, b, c), T(x_1; a, b, c)]\, \chi_{x_1 \le b} \\ & + [T(x_1; a, b, c), T(x_0; a, b, c)]\, \chi_{b \le x_0} \\ & + [\min\{T(x_0; a, b, c), T(x_1; a, b, c)\}, 1]\chi_{x_0 < b < x_1} \end{aligned} \tag{10}$$

Thus, the natural extension of (8) to interval fuzzy values of the evidence is

$$\Box\widetilde{L}(H_j|\,[E]) \propto \sum\nolimits_{e_L \in \mathcal{L}(\mathcal{E})} \Box\widetilde{e_L}([E])p(e_L|H_j)$$

i.e. an interval $\left[\left(\Box\widetilde{L}(H_j|\,[E])\right)_0, \left(\Box\widetilde{L}(H_j|\,[E])\right)_1\right]$. If $[E] = [E_0, E_1]$ then $\left(\Box\widetilde{L}(H_j|\,[E])\right)_0 = \Box\widetilde{L}(H_j|E_0)$ and $\left(\Box\widetilde{L}(H_j|\,[E])\right)_1 = \Box\widetilde{L}(H_j|E_1)$, calculated according to (8).

The lower and upper likelihoods provides a lower/upper class of inclusion functions for the unnormalized posteriors,

$$\Im^* = \{\Box\widetilde{p}^* : \forall H_j, \Box\widetilde{l}^*\,(H_j|\,[E]) \le \Box\widetilde{p}^*\,(H_j|\,[E]) \le \Box\widetilde{u}^*\,(H_j|\,[E])\} \tag{11}$$

where $\Box\widetilde{l}^*\,(H_j|\,[E])$ and $\Box\widetilde{u}^*\,(H_j|\,[E])$ are proportional to $\left(\Box\widetilde{L}(H_j|\,[E])\right)_0 p(H_j)$ and to $\left(\Box\widetilde{L}(H_j|\,[E])\right)_1 p(H_j)$, respectively.

Let $\Im$ be the class of normalized inclusion functions $\Box\widetilde{p}(H_j|\,[E])$ for the posteriors. If we whish a lower/upper expression for the class $\Im$, we have to minimize and maximize $\Box\widetilde{p}(H_j|\,[E])$ over $\Im^*$. To minimize it, we take

$$\Box\widetilde{p}^*(H|\,[E]) = \Box\widetilde{l}^*(H|\,[E])\chi_{H=H_j} + \Box\widetilde{u}^*(H|\,[E])\chi_{H\neq H_j} \tag{12}$$

in order to assign minimum possible mass to H_j and maximal mass to every $H \neq H_j$. Thus, we obtain the *normalized lower posterior*,

$$\Box\widetilde{l}\,(H_j|\,[E]) = \Box\widetilde{l}^*(H_j|\,[E]) \Big/ \left(\Box\widetilde{l}^*(H_j|\,[E]) + \sum\nolimits_{\substack{k=1\\k\neq j}}^{m} \Box\widetilde{u}^*(H_k|\,[E])\right) \tag{13}$$

To maximize $_\square\widetilde{p}(H_j|\,[E])$ simply swap $_\square\widetilde{l}^*$ with $_\square\widetilde{u}^*$ in (12). Swapping $_\square\widetilde{l}^*$ with $_\square\widetilde{u}^*$ in (13) we obtain the expresion for the *normalized upper posterior* $_\square\widetilde{u}\,(H_j|\,[E])$.

The following proposition shows that although neither $_\square\widetilde{l}$ nor $_\square\widetilde{u}$ are additive over Θ they both distribute a whole mass of two between any subset $A \subseteq \Theta$ and its complement $\Theta\backslash A$.

Proposition 1 *The lower/upper posteriors satisfy:*

(i) $1 \in \mathrm{core}_\square\widetilde{l}(\Theta|\,[E]) \cap \mathrm{core}_\square\widetilde{u}(\Theta|\,[E])$*;*

(ii) $_\square\widetilde{l}\,(A|\,[E]) + {_\square\widetilde{l}}((\Theta\backslash A)|\,[E]) \preccurlyeq \mathsf{one} \preccurlyeq {_\square\widetilde{u}}\,(A|\,[E]) + {_\square\widetilde{u}}((\Theta\backslash A)|\,[E])$*;*

(iii) $_\square\widetilde{l}\,(A|\,[E]) + {_\square\widetilde{l}}((\Theta\backslash A)|\,[E]) + {_\square\widetilde{u}}\,(A|\,[E]) + {_\square\widetilde{u}}((\Theta\backslash A)|\,[E]) = \mathsf{two}$.

Proof. (i) Let $\widetilde{l}^*_j$, $\widetilde{u}^*_j$, $\widetilde{l}^*_\theta$ and $\widetilde{u}^*_\theta$ denote $_\square\widetilde{l}^*(H_j|\,[E])$, $_\square\widetilde{u}^*(H_j|\,[E])$, $_\square\widetilde{l}^*(\Theta|\,[E])$ and $_\square\widetilde{u}^*(\Theta|\,[E])$, respectively. For every $\alpha \in (0,1]$, the α-cut of $_\square\widetilde{l}$ is

$$ {}^\alpha{}_\square\widetilde{l}(H_j|\,[E]) \quad = \quad \left[\frac{\left(\widetilde{l}^*_j\right)^\alpha_0}{\left(\widetilde{l}^*_j\right)^\alpha_1 + \sum_{\substack{k=1\\k\neq j}}^m \left(\widetilde{u}^*_k\right)^\alpha_1}, \frac{\left(\widetilde{l}^*_j\right)^\alpha_1}{\left(\widetilde{l}^*_j\right)^\alpha_0 + \sum_{\substack{k=1\\k\neq j}}^m \left(\widetilde{u}^*_k\right)^\alpha_0} \right] \tag{14} $$

Swapping $\widetilde{l}^*_\bullet$ with $\widetilde{u}^*_\bullet$ we obtain the expression for the α-cut of $_\square\widetilde{u}$.

Let X be a fuzzy subset, and let X^α_0 and X^α_1 be the left and right endpoints of its α-cut. Then

$$ {}^\alpha{}_\square\widetilde{l}(\Theta|\,[E]) = \left[\left(\widetilde{l}^*_\theta\right)^\alpha_0 \Big/ \left(\widetilde{l}^*_\theta\right)^\alpha_1, \left(\widetilde{l}^*_\theta\right)^\alpha_1 \Big/ \left(\widetilde{l}^*_\theta\right)^\alpha_0 \right] \tag{15} $$

because the summations are over empty ranges. On the other hand, for every fuzzy subset X and $\forall\alpha \in (0,1]$, $X^\alpha_0 \leq X^\alpha_1$, and then $\forall\alpha \in (0,1]$, $1 \in {}^\alpha\widetilde{l}(\Theta|\,[E])$; i.e. $1 \in \mathrm{core}_\square\widetilde{l}(\Theta|\,[E])$. We can prove that $1 \in \mathrm{core}_\square\widetilde{u}(\Theta|\,[E])$ in a similar way.

(ii) Let $C \subseteq \Theta$. Let $\widetilde{l}^*_C$ and $\widetilde{u}^*_C$ denote $_\square\widetilde{l}^*(C|\,[E])$ and $_\square\widetilde{u}^*(C|\,[E])$, respectively. Then

$$ \begin{aligned} & {_\square\widetilde{l}}\,(A|\,[E]) + {_\square\widetilde{l}}((\Theta\backslash A)|\,[E]) \\ = \; & \frac{_\square\widetilde{l}^*(A|\,[E])}{_\square\widetilde{l}^*(A|\,[E]) + {_\square\widetilde{u}^*}((\Theta\backslash A)|\,[E])} + \frac{_\square\widetilde{l}^*((\Theta\backslash A)|\,[E])}{_\square\widetilde{l}^*((\Theta\backslash A)|\,[E]) + {_\square\widetilde{u}^*}(A|\,[E])} \\ = \; & (\widetilde{l}^*_A\widetilde{u}^*_A + 2\widetilde{l}^*_A\widetilde{l}^*_{\Theta\backslash A} + \widetilde{l}^*_{\Theta\backslash A}\widetilde{u}^*_{\Theta\backslash A}) \Big/ (\widetilde{l}^*_A\widetilde{u}^*_A + \widetilde{l}^*_A\widetilde{l}^*_{\Theta\backslash A} + \widetilde{l}^*_{\Theta\backslash A}\widetilde{u}^*_{\Theta\backslash A} + \widetilde{u}^*_A\widetilde{u}^*_{\Theta\backslash A}) \\ = \; & (\widetilde{k} + \widetilde{l}^*_A\widetilde{l}^*_{\Theta\backslash A}) \Big/ (\widetilde{k} + \widetilde{u}^*_A\widetilde{u}^*_{\Theta\backslash A}) \end{aligned} \tag{16} $$

where $\widetilde{k}$ is a fuzzy constant. As $\widetilde{l}^*_C \preccurlyeq \widetilde{u}^*_C$ we obtain $_\square\widetilde{l}\,(A|\,[E]) + {_\square\widetilde{l}}((\Theta\backslash A)|\,[E]) \preccurlyeq \mathsf{one}$. The another inequality in (ii) is proved in a similar way.

(iii) It is enough to add the following and (16) together

$$ {_\square\widetilde{u}}\,(A|\,[E]) + {_\square\widetilde{u}}((\Theta\backslash A)|\,[E]) = \frac{\widetilde{l}^*_A\widetilde{u}^*_A + 2\widetilde{u}^*_A\widetilde{u}^*_{\Theta\backslash A} + \widetilde{l}^*_{\Theta\backslash A}\widetilde{u}^*_{\Theta\backslash A}}{\widetilde{l}^*_A\widetilde{u}^*_A + \widetilde{l}^*_A\widetilde{l}^*_{\Theta\backslash A} + \widetilde{l}^*_{\Theta\backslash A}\widetilde{u}^*_{\Theta\backslash A} + \widetilde{u}^*_A\widetilde{u}^*_{\Theta\backslash A}} $$

■

0.5 An Example In Health Care

Suppose that we have collected sound level data at a place under study. Also we have interviewed population about their annoyance response to noise at this place. Assume that some time in the future, we are interested in knowing, only from sound level measurements, without the need of new surveys, if people living at this place, is annoyed because of acoustic contamination.

Table 0.1: *Fuzzy t-numbers in describing interval fuzzy and interval probabilities*

TFN_5	a	b	c	TFN_7	a	b	c
very low	$-1/4$	0	$1/4$	zero	$-1/6$	0	$1/6$
low	0	$1/4$	$1/2$	very low	0	$1/6$	$1/3$
medium	$1/4$	$1/2$	$3/4$	low	$1/6$	$1/3$	$1/2$
high	$1/2$	$3/4$	1	medium	$1/3$	$1/2$	$2/3$
very high	$3/4$	1	$5/4$	high	$1/2$	$2/3$	$5/6$
				very high	$2/3$	$5/6$	1
				one	$5/6$	1	$7/6$

In estimating annoyance level, we use a fuzzy opted quantization into five exhaustive and mutually exclusive states: absent, mild, moderate, intense, or severe. Given sound level measurements from a place, we wish to classify people's annoyance response to noise at this place into one of these five situations according to these measurements and previous knowledge.

Let T_1 and T_2 denote the triangular fuzzy numbers (a_1, b_1, c_1) and (a_2, b_2, c_2). The basic operations wich we use are:

$$T_1 \overset{+}{\times} T_2 = (a_1 \overset{+}{\times} a_2, b_1 \overset{+}{\times} b_2, c_1 \overset{+}{\times} c_2) \tag{17}$$
$$-T_1 = (-c_1, -b_1, -a_1) \tag{18}$$
$$1 \diagup T_1 \cong (1/c_1, 1/b_1, 1/a_1) \tag{19}$$

where $\cong$ denotes approximation [6]. According to Zadeh's extension principle, the extension of the tfn $T = (a, b, c)$ to the fuzzy sets is given by[1]

$$\widetilde{T}(A; a, b, c)(y) = \begin{cases} 0 & \text{if } 1 < y \\ \max\left\{A\left(a + (b-a)y\right), A\left(c + (b-c)y\right)\right\} & \text{if } y \leq 1 \end{cases} \tag{20}$$

Thus, we naturally extend (10) to

$$\begin{aligned} \Box\widetilde{T}([E]\,; a, b, c) = & \left[\widetilde{T}(E_0; a, b, c), \widetilde{T}(E_1; a, b, c)\right] \chi_{x_1 \leq b} && (21) \\ & + \left[\widetilde{T}(E_1; a, b, c), \widetilde{T}(E_0; a, b, c)\right] \chi_{b \leq x_0} \\ & + [\min\{\widetilde{T}(E_0; a, b, c), \widetilde{T}(E_1; a, b, c)\}, 1]\chi_{x_0 < b < x_1} && (22) \end{aligned}$$

We assume that sound level measurements are reported according to an interval with triangular fuzzy numbers as endpoints. These "talking" sound level meters use the term-set $\text{TFN}_0^{10} = \{$zero, one, ..., ten$\}$. Because of the subjective nature of annoyance, evaluation must be carried out using survey techniques such as questionnaires. The term-set $\text{TFN}_5 = \{$very low, low, medium, high, very high$\}$ (Cf. Table 0.1) is used by experts and by the general population to classify sound level average regarding to a place. Similarly, in describing interval fuzzy probabilities for the final conclusions we consider the term-set $\text{TFN}_7 = \{$zero, very low, low, medium, high, very high, one$\}$ (Cf. Table 0.1). Sound level average linguistic estimation and probabilities could be represented by any ordered term-pair from TFN_5 and from TFN_7, respectively.

"Data-based" priors for any annoyance level state could be deduced from information contained in samples of past experiences in which the same noise was studied. Assume that the prior distribution and the likelihood probabilities are given as shown in Tables 0.2 and 0.3, respectively. For example, a zero-valued likelihood probability p(very low | severe) is reasonable, because it means that zero is the probability —regarded as representing a degree of reasonable belief or confidence rather than a frequency— that an expert classifies as very low the sound level average regarding to a place where population's annoyance is severe (Cf. Table 0.3).

[1] Let $\mathcal{U}$ and $\mathcal{V}$ be two spaces, $f : \mathcal{U} \to \mathcal{V}$ a function, and $S \in \mathfrak{F}(\mathcal{U})$. The image of S by the extension of f is $S' \in \mathfrak{F}(\mathcal{V})$, constructively defined as: $\forall y \in \mathcal{V}, S'(y) = [\sup\{S(x) | x \in \mathcal{U} \wedge y = f(x)\}$ if $f^{-1}(y) \neq \varnothing$; 0 else]. If S is $T(x; a, b, c) = \max(\min((x-a)/(b-a), 1, (c-x)/(c-b)), 0)$ then S' is $\widetilde{T}(A; a, b, c)$ given by (20).

Table 0.2: *Annoyance priors*

Population's annoyance	$p(\cdot)$
absent	very low
mild	very low
moderate	high
intense	zero
severe	zero

Table 0.3: *Likelihoods*

Population's annoyance	very low	low	medium	high	very high
absent	very high	very low	zero	zero	zero
mild	very low	low	low	very low	zero
moderate	zero	very low	high	very low	zero
intense	zero	zero	zero	high	low
severe	zero	zero	zero	zero	one

We are interesting in the maximum a posteriori (MAP) estimation. For each interval posterior $[_\square\widetilde{l}\,(H_j|\,[E])\,,_\square\widetilde{u}\,(H_j|\,[E])]$ we calculate the mean fuzzy set of its fuzzy endpoints. Baas and Kwakernaak's fuzzy number ranking procedure [7] allows us to find the MAP. The final step is to find the nearest words —belonging to TFV$_7$— to the fuzzy endpoints of the MAP. This is accomplished by using a distance defined on triangular fuzzy numbers. Given $T_1 \equiv (a_1, b_1, c_1)$ and $T_2 \equiv (a_2, b_2, c_2)$, we use an unweighted Euclidean format [8]: $d(T_1, T_2) = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2}$.

For the data showed in Tables 0.1-0.3, and for an interval fuzzy measurement —reported by a "talking" sound level meter— of [six, seven], then, with an interval probability of [high, very high] the best alternative is that people living at that place is moderate annoyed because of acoustic contamination.

0.6 Conclusions

We have centered on how to learn under a Bayesian point of view from imprecise linguistic data (evidence). We have assumed that linguistic imprecision lies at the same time in evidence, priors and marginals, and accordingly in the posteriors. We represent this linguistic imprecision using interval fuzzy and interval probability assignments. Our linguistic imprecision propagation Bayesian mechanism allows us to express conclusions in a natural way such as: "In the light of these results we can say that for a sound level of approximately six or seven, then, with a probability, approximately high or perhaps very high, the best alternative is that people living at that place is moderate annoyed because of acoustic contamination".

References

[1] Walley, P., Gurrin, L. and Burton, P. (1996). Analysis of clinical data using imprecise prior probabilities. *The Statistician* **45**(4), 457–485. 0.2, 0.2

[2] DeRobertis, L. and Hartigan, J. A. (1981). Bayesian inference using intervals of measures. *Annals of Statistics* **9**, 235–244. 0.2

[3] Nakoula, Y., Galichet, S. and Foulloy, L. (1997). Identification of linguistic fuzzy models based on learning. In: Hellendoorn, H. and Driankov, D. (eds.) *Fuzzy Model Identification. Selected Approaches.* Springer-Verlag, Berlin Heidelberg New York, 281–319. 0.3

[4] Itoh, S. and Itagaki, H. (1989). Applications of fuzzy-bayesian analysis to structural reliability. In: *Proceedings of ICOSSAR'89, the 5th International Conference on Structural Safety and Reliability.* San Francisco. 0.3

[5] Yang, C.C. (1997). Fuzzy bayesian inference. In: *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics.* Orlando, 2707–2712. 0.3

[6] Triantaphyllou, E. and Lin, Ch.-T. (1996). Development and evaluation of five fuzzy multiattribute decision-making methods. *International Journal of Approximate Reasoning* **14**, 281–310. 0.5

[7] Baas, S.J. and Kwakernaak, H. (1977). Rating and ranking of multi-aspect alternatives using fuzzy sets. *Automatica* **13**, 47–58. 0.5

[8] Diamond, P. (1989). Fuzzy kriging. *Fuzzy Sets and Systems* **33**, 315–332. 0.5